ELSEVIER

Journal of Chromatography A, 972 (2002) 27–43

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Intelligent automation of high-performance liquid chromatography method development by means of a real-time knowledge-based approach

Ting-Po I[a], Randy Smith[a], Sam Guhan[a], Ken Taksen[a], Mark Vavra[b], Douglas Myers[c], Milton T.W. Hearn[d,*]

[a]*PGRD Pfizer Inc., Eastern Point Road, Groton, CT 06340, USA*
[b]*Gas Technology Institute, 1700 S. Mount Prospect Road, Des Plaines, IL 60018, USA*
[c]*Intelligent Laboratory Solutions, Inc., 416 Braemar Avenue, Naperville, IL 60563, USA*
[d]*Centre for Bioprocess Technology & Australian Centre for Research on Separation Science, Monash University, Clayton 3800, Australia*

Abstract

We describe the development, attributes and capabilities of a novel type of artificial intelligence system, called LabExpert, for automation of HPLC method development. Unlike other computerised method development systems, LabExpert operates in real-time, using an artificial intelligence system and design engine to provide experimental decision outcomes relevant to the optimisation of complex separations as well as the control of the instrumentation, column selection, mobile phase choice and other experimental parameters. LabExpert manages every input parameter to a HPLC data station and evaluates each output parameter of the HPLC data station in real-time as part of its decision process. Based on a combination of inherent and user-defined evaluation criteria, the artificial intelligence system programs use a reasoning process, applying chromatographic principles and acquired experimental observations to iteratively provide a regime for a priori development of an acceptable HPLC separation method. Because remote monitoring and control are also functions of LabExpert, the system allows full-time utilisation of analytical instrumentation and associated laboratory resources. Based on our experience with LabExpert with a wide range of analyte mixtures, this artificial intelligence system consistently identified in a similar or faster time-frame preferred sets of analytical conditions that are equal in resolution, efficiency and throughput to those empirically determined by highly experienced chromatographic scientists. An illustrative example, demonstrating the potential of LabExpert in the process of method development of drug substances, is provided.

Keywords: Method development; Expert systems; Automation

1. Introduction

High-performance liquid chromatography (HPLC) has become a universal tool for pharmaceutical and biomedical research, as well as for product analysis. HPLC applications in their various selectivity modes now include the separation and analysis of new compounds, determination of sample purity, detection and isolation of impurities, and purification of the final product. Within the context of pharmaceutical analysis, automation of method development

*Corresponding author. Tel.: +61-3-990-53720; fax: +61-3-9905-5882.
E-mail address: milton.hearn@med.monash.edu.au (M.T.W. Hearn).

0021-9673/02/$ – see front matter
PII: S0021-9673(02)01075-0

becomes an important objective, avoiding time-consuming and repetitive tasks and thus providing substantial practical benefits. Notwithstanding the significant progress that has been achieved over the past decade, optimisation of a specific HPLC separation of complex mixtures of analytes, particularly those from biological sources, can still require a great deal of specialised knowledge. The complexity of the task to optimise the parameters of interest in chromatography (and capillary electrophoresis) has been recognised for decades [1,2]. The analyst must evaluate the problem and take into account a large number of system parameters and variables, including the nature of the solute–solvent inter-actions, adsorption isotherms, temperature, column dimensions, packing characteristics, particle and pore size of the sorbent, the phase ratio and chemical characteristics of the sorbent as well as the flow-rate. While the influence of some of these parameters and variables can be assessed or predicted by chromatographic theory, effective method development still requires extensive collection, interpretation, prioritisation and simulation of experimental data, i.e., application of the knowledge-based experience of the human end-user.

Many of the individual aspects associated with method development processes in analytical chemistry have been presented in the literature. Typical of these strategies, the experimental path for HPLC method development and automated solid-phase extraction for the analysis of polar and non-polar pharmaceutical substances has been extensively discussed in the literature [3–6], as have the experimental methodologies for HPLC method development for chiral compounds and other complex mixtures [7]. Independently, computer-assisted HPLC method optimisation schemes involving partial automation of individual chromatographic tasks [8–10], as expert systems for the fault diagnosis in gas chromatography (GC) or in GC–mass spectroscopic system identifications [11], or use of simplex algorithms with diode array detection for optimisation of HPLC separations [12] have also been described. Other approaches to HPLC method development have included decision criteria based on compound chemical structures or databases of column selectivity variables, coupled to the use of computer-assisted predictions such as EluEx [13,14], ChromSword RTM [15] or other algorithms to optimise existing HPLC separation methods [16], and methods for modelling and optimising the resolution response characteristics of the chromatographic system [17–19].

In this paper, we describe a new concept in machine-based expert systems that applies an artificial intelligence program, which we have called LabExpert, to automate a substantial part of the data collection, interpretation and optimisation in HPLC method development. The benefits of this new approach include substantial timesaving and more effective use of laboratory facilities, staff and resources. In addition, by incorporating and assimilating the reasoning processes of expert chromatographers, LabExpert makes their knowledge and experience available to less well-trained or experienced chromatographic scientists, increasing laboratory productivity and aiding in their productivity. Because the strategy followed is not dependent on the type of HPLC sorbent or the nature of the elution conditions employed, application of the concepts and procedures described in this manuscript will be of general utility to most fields of analytical high-performance separation of polar and neutral analytes, including orthogonal techniques such as capillary liquid chromatography (μLC), capillary electrophoresis (cap-HPCE) and capillary electrochromatography (CEC), as well as hyphenated experimental procedures such as capillary LC–MS and CE–MS.

2. Experimental

2.1. General considerations

The experimental procedures described below represent a typical separation study carried out in our laboratories leading to the development of LabExpert and illustrate the decision processes that the system incorporates as it proceeds from an initial unresolved chromatogram to a criterion-specified, acceptable separation. The acceptability criteria included achievement of minimum defined values of resolution (R_s), selectivity (α), column plate number (N), peak asymmetry factor (λ) and analysis time (t_{total})

for a specified isocratic separation. The results described below represent an actual isocratic method development investigation, in this case involving the separation of four proprietary Pfizer compounds, each of which is identified by a code number. Associated investigations to be reported subsequently relate to the application of these concepts with gradient elution systems. Knowledge of the actual structure or chemical functionality of the compounds is not a prerequisite to method development using LabExpert, although the software has been designed to incorporate molecular descriptors, such as pK_a, intrinsic hydrophobicity, dipole moment, net charge, other physico-chemical parameters and molecular coefficient indices and molecular structure, as part of the optimisation subroutines when prediction of retention behaviour is planned. The LabExpert software was written around the framework of the G2 computer language platform (Gensym, Burlington, MA, USA), which is an industrial real-time expert system development language, suited for creating software applications that applies decision-making and reasoning about objects in real time [20]. In contrast to traditional procedural programming languages such as Fortran or C/C++, G2 utilises complex objects to model the dynamic properties and behaviours of applications. By providing powerful data system, device connectivity and real-time reasoning abilities, G2 was the language of choice for incorporating chromatographic knowledge into a machine-based expert system.

2.2. Instrumentation

Data were acquired using an Agilent (HP) 1100 HPLC system including a diode array detector and quaternary pump system, equipped with degasser, autosampler and Chemstation software. A LC Spiderling 8 position column changer from Chiralizer Services (Newtown, PA, USA) was mounted with the following HPLC columns: Zorbax SB-CN (150×4.6 mm, 5 μm, Agilent, Littlefalls, DE, USA), Waters Symmetry C_{18} (150×4.6 mm, 5 μm, Waters, Milford, MA, USA), Inertsil C_8 (150×4.6 mm, 5 μm, Phenomenex, Torrance, CA, USA). Expert chromatographic knowledge was gathered from ''knowledge engineering'' sessions with a team of experienced analytical chemists/pharmaceutical scientists working at Pfizer Inc. and other academic scientists, with LabExpert incorporating sets of relevant operationally defined rules and procedures, heuristic guidelines and appropriate equations derived from linear and non-linear elution chromatographic theory. The LabExpert software running on G2 version 5.0 Rev.3 was installed on Compaq Deskpro Pentium II computers and interfaced to Agilent (Hewlett-Packard) Chemstations through a G2-ActiveXLink software bridge.

The four solvent reservoirs of the Agilent series 1100 HPLC system were filled with the following mobile phases: (A) 25 m*M* KH_2PO_4, pH 3.5; (B) acetonitrile (ACN); (C) methanol; (D) water. The detector was set at 230 nm, column temperature was ambient, and the injection volume was 20 μl. Samples identified as CP-547,224, CP-195,543, CP-316,698, and CP-547,238 are proprietary Pfizer medicinal compounds and were prepared for analysis by weighing 12.5 mg of each compound into individual 25 ml volumetric flasks, dissolving with sonication, and bringing to volume with acetonitrile. A 5-ml aliquot of each solution was transferred by pipette into individual 25-ml volumetric flasks, brought to volume with acetonitrile and then mixed thoroughly. The sample injection solution consisted of a mixture of equal parts of these analytes diluted to appropriate concentrations in acetonitrile.

3. Results and discussion

3.1. Incorporation of artificial intelligence systems into the analytical chemistry domain

The field of artificial intelligence (AI) came into existence in the 1950s [21] with the objective to make computers more intelligent and to permit machines to emulate human capabilities. The derived research area called ''Expert Systems'' has since demonstrated a broad range of capabilities associated with difficult problem-solving tasks in areas of planning, design, process control, scheduling and diagnosis [11,22]. An expert system can be defined as ''a computer program based on artificial intelli-

gence techniques, performing a specialised difficult task at the level of a human expert'' [23]. Induction procedures based on information theory/probability algorithms and on self-organising neural network methods have been used to classify and direct outcome from machine learning and artificial intelligence [24]. Expert systems have been particularly successfully when the scope of the problem is fixed and relatively narrow. For example, a focused expert system used in chemical reactor design would probably not be useful for designing automobile engines.

Expert systems have proven very successful in applications where a complex series of possible solutions result in a large ''search space''. Human expertise is often incorporated into such systems, usually as heuristic rules, to constrain the problem and to limit the search space. In the context of analytical chemistry, a number of strategies have gained favour for the application of expert system technologies, including implementation of the ACExpert project for GC–MS analysis of trace organics in environmentally significant samples [11,25]. Similarly, the NINA software [26] and related algorithms have been used (a) for method setup and optimisation [27] in ion chromatography; (b) in trace analysis of pesticide residues and other environmental contaminants [28]; (c) in performance assessment and instrumentation diagnosis in gas chromatography with standard mixtures [29]; (d) for diagnosing the cause of problematic atomic absorption spectrometry data [30]; (e) for amino acid sequence identification [31]; and (f) for engineering representations of method validation and trouble shooting analysis [32,33]. A further powerful attribute of expert systems is their potential for accurate prediction of distribution coefficients and retention indices in gas and liquid chromatography from knowledge of the structure of the analytes [13–15,34,35]. In various cases, chemometric methods using a series of statistical tests, principal component regression (PCR), linear discriminant analysis (LDA) or other pattern recognition algorithms have also been included [36,37]. Often these chemometric procedures have been associated with multiple classification ripple-down rules [38], domain-oriented knowledge acquisition tools [39] or inference engines [37], such as the EAengine, to enable classification of experimental data, to offer conclusions on the attributes of a particular system or to enable subsequent prediction of the attributes of unknown sets of conditions or system features. In the evolution of the LabExpert system, this background has been drawn upon, but importantly LabExpert diverges in a number of significant and fundamental ways, both in terms of its architectural logic and also in the manner that the expert rules are initially coded and incorporated into the software as part of a machine-learning strategy that is subsequently followed.

3.2. Evolution of HPLC control systems

LabExpert represents an approach to a new generation of expert systems for the design, supervisory control, analysis and execution of HPLC method development. The evolution of LabExpert system has been enabled by evolution in the design philosophy of manufacturers of HPLC instrumentation towards more open architecture and increased software control of the HPLC hardware. For optimal performance, communication with a real-time expert system such as LabExpert requires a single point, interactive control of the entire HPLC equipment system. Until recently, HPLC automation and control have incorporated dedicated proprietary computers and software programs as ''closed systems'' that did not allow any user program control. Arising from the present investigations, the LabExpert system achieves a level of ''omniscience'' through the software's bi-directional control link to HPLC instrumentation. This bi-directional configuration permits the simultaneous operation of several HPLC (or HPCE, μLC–MS and HPCE–MS) instruments coupled via TCP/IP connections via appropriate interfaces to a specific LabExpert application. Chromatographic knowledge based on heuristic rules of various experts has been incorporated into the LabExpert software, together with the knowledge of the instrumental format. The decision trees in LabExpert contain or can directly access all of this chromatographic knowledge. In the LabExpert format, the machine-learning expert system has been programmed to follow the same principles and decision processes as used by expert chromatographers, designing a series of HPLC experiments to obtain an

acceptable separation for a given sample set. The G2-Bridge layer provides automation as well as an interface to the HPLC equipment. The language specific to each item of equipment is provided at the G2-Bridge level.

3.3. The chromatographic scientist as an expert system

As a ''user-friendly'' starting point, the techniques intuitively used [40] by experienced chromatographic scientists can be applied to inaugurate the initial machine learning applications of the LabExpert system. LabExpert has been programmed to use intelligent automation to enhance laboratory productivity by emulating the expert chromatographer's knowledge-based activities of method design and results interpretation. The key factor that makes HPLC method development an ideal candidate for such control by machine-based expert systems is that it consists of routines that can be broken into subroutines and linked together by a series of decision trees. Irrespective of whether inductive, empirical or chemometric principles are then applied, acceptable separations are achieved when the analyst's criteria are meet for a designated resolution (R_s), normalised retention time (k') and selectivity (α). Traditionally, to reach this outcome, the experience of the analyst is combined with his/her familiarity with the operation of the HPLC system and fundamental knowledge of chromatographic theory to develop a separation. Various empirical matrix searching techniques, such as DryLab 2000 or PESOS, have previously been used to streamline the selection of these experiments using two or more reasonable separations. However, these programs do not control or physically automate the independent operation of the HPLC system, select initial starting conditions for the separations or make and execute decisions as an expert system. In comparison, the LabExpert system, in common with the interpretative and decision processes followed by experienced chromatographers, provides real-time interpretation of the results, comparing them to pre-defined acceptability criteria with the ability to reason about previous results and implements on-line changes in a fully automated manner.

3.4. Design elements underpinning the use of labExpert system in method development

The development of a HPLC separation method involves an iterative process of planning, method implementation, execution and interpretation. Traditionally, the analyst first establishes a general plan for developing an initial set of analytical conditions based on chromatographic theory in conjunction with prior knowledge and experience. A method, derived from this plan, is then implemented and includes the choice of a column and mobile phase (isocratic or gradient) composition. Execution of the plan provides the chromatographer with additional data, which, after interpretation, suggests what adjustments need to be implemented as a sequence of further decision and action steps. The ideal method development obviously involves a minimum number of carefully planned and executed HPLC experiments to obtain an acceptable separation.

Current laboratory automation involves using computer-controlled, robotic HPLC systems to physically execute this sequence of unit operations, which includes setting flow-rates, controlling pumps, manipulating valves, injecting samples and data collection. This level of automation allows the system to run unattended and eliminates many of the mundane manual operations. A chromatographer is able to program a sequence of methods using menu-driven routines offered within the HPLC software, permitting unattended operation of these sequences. Although automation of the physical injection and column re-equilibration sequence eliminates a substantial part of the routine operation, a significant investment of the analyst's time still involves the routine planning of the next sequence of experiments leading to improvement in the quality of the data collection. This planning includes consideration of the column type, the mobile phase composition and sorbent particle size or porosity. The chromatographer incorporates his/her expertise at the beginning of this Grid-Search to determine an appropriate set of selectivity variables for the samples to be run, and then again at the end of the sequence to interpret and iteratively refine the results and to continue with the most likely method identified by that sequence. In this mode, continuous input from the analyst is a vital and necessary part of the method development

loop. A primary objective of all artificial intelligence systems in HPLC analysis is to at least achieve resolution outcomes that match in real-time the decisions and criteria of the expert chromatographer. With the development of LabExpert, this loop is closed with regard to the planning steps required for machine-based intelligent reasoning about the latest results, thus facilitating better time utilisation and resource management.

3.5. Computer-assisted search routines

Several computerised programs based on Grid-Search software [16,41] have been developed in an attempt to automate chromatographic optimisation decision processes. Techniques utilising Grid-Search software perform a large number of experiments, usually varying only two parameters or variables. Results of these experiments, which are independently planned and executed, are correlated to one another, using chromatographic theory to predict an optimal set of analytical conditions. These Grid-Search procedures, however, require a substantial time investment and often generate redundant or irrelevant results since they have no mechanism to decide what regions of the separation space will be less or more productive. Other methods based on simplex algorithms [42–44] or factorial design [34,45] routines, though useful, have the disadvantage of also requiring a large number of experiments to reach a useful solution to the separation problem. Moreover, these procedures focus on the planning, rather than outcome, stages of the method development process. Predictions based on chromatographic theory or on the elution behaviour of test analytes can be incorporated into the initial stages of the Grid-Search-based process of method development process. This type of optimisation yields results that then need further refinement. Grid searches enacted on integrated HPLC workstations or equipment may provide insight into the solution for a particular separation problem, but they require a substantial investment of materials and time. As currently practiced, with Grid-Search procedures a chromatographer's active interpretation and involvement must accompany method discovery and moreover such computer-aided techniques do not select initial starting conditions for the separations or make and execute decisions related to the operational automation and control of the HPLC system. In a true expert system, chromatographic knowledge is incorporated into the system as part of the machine-learning tasks allowing real-time chromatographic control.

3.6. LabExpert system for HPLC: real-time artificial intelligence

In traditional approaches to system optimisation, expert decisions and input are needed at each step of the method development process, involving interpretation of the analysis by a knowledgeable scientist. Conditions for subsequent analysis incorporate decisions based on this interpretation. When the sequence is complete, the chromatographer examines the results to assess whether adequate separation has been achieved or whether another experimental sequence is required. This process has now been closely modelled by the computerised artificial intelligence system, LabExpert. The essential features of LabExpert system are its real-time capabilities, dynamic virtual object architecture, planning engine and interconnected decision trees that permit autonomous operation, system surveillance, on-line data evaluation and intelligent automation of HPLC method development.

The architecture of LabExpert contains a number of software objects that represent, in a virtual context, individual items of the HPLC equipment system, such as samples, columns and solvent containers. These objects have attributes describing their configuration in the system and have associated behaviours. There are also user-preference objects that allow the user to select how evaluation criteria can be applied to the method analyses. Based on initial input information from the operator, LabExpert creates an initial plan to accomplish a sequence of tasks such as a gradient scouting method to identify whether a column or a mobile phase choice has the desired separation characteristics for a particular sample. LabExpert then instructs the HPLC equipment to run the HPLC analysis and monitors the instrument for operating parameters, such as pump pressure, solvent availability, flow-rate and temperature. During or at the conclusion of the analytical run, LabExpert up-loads information on the chromatographic profile, such as peak area, width, height, symmetry, and spectral purity from the HPLC workstation. LabExpert incorporates chro-

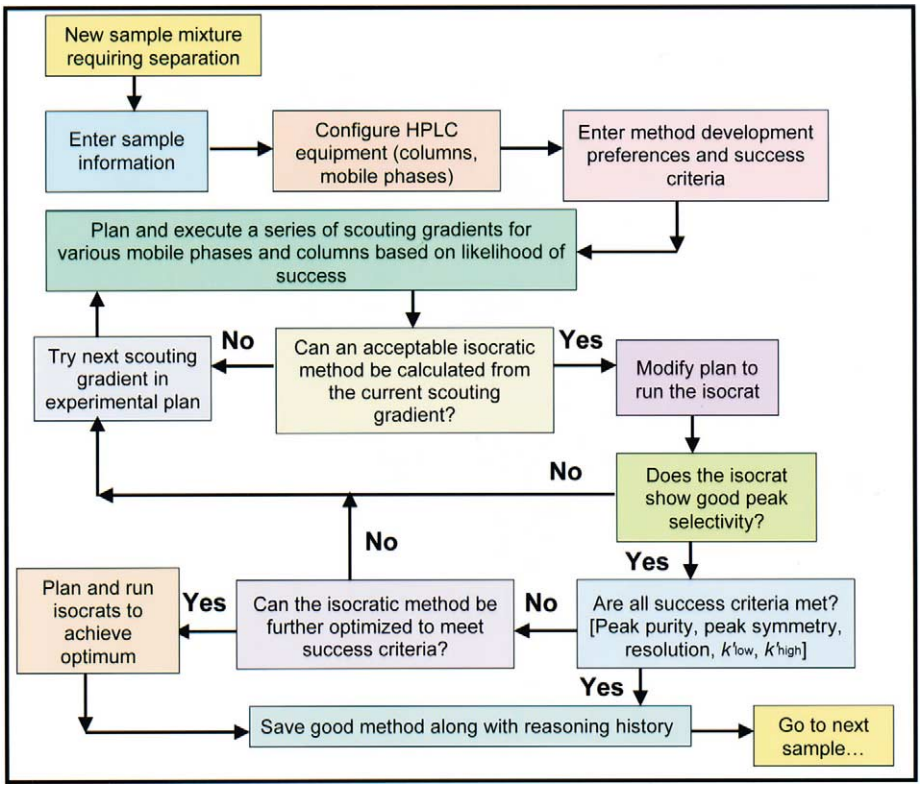


Fig. 1. The main logic flow involved in the design and operation of the decision trees implemented within LabExpert framework. Starting with the mixture requiring separation (yellow box), the user provides inputs of sample information, equipment configuration, and success criteria (colour boxes, second row) into the LabExpert software. The software then creates and executes a plan controlled by various decisions (colour boxes, third to seventh row) based on the use of the expert rules leading to the development of an acceptable method. The results and the method development history are stored and can be printed out in a summary report.

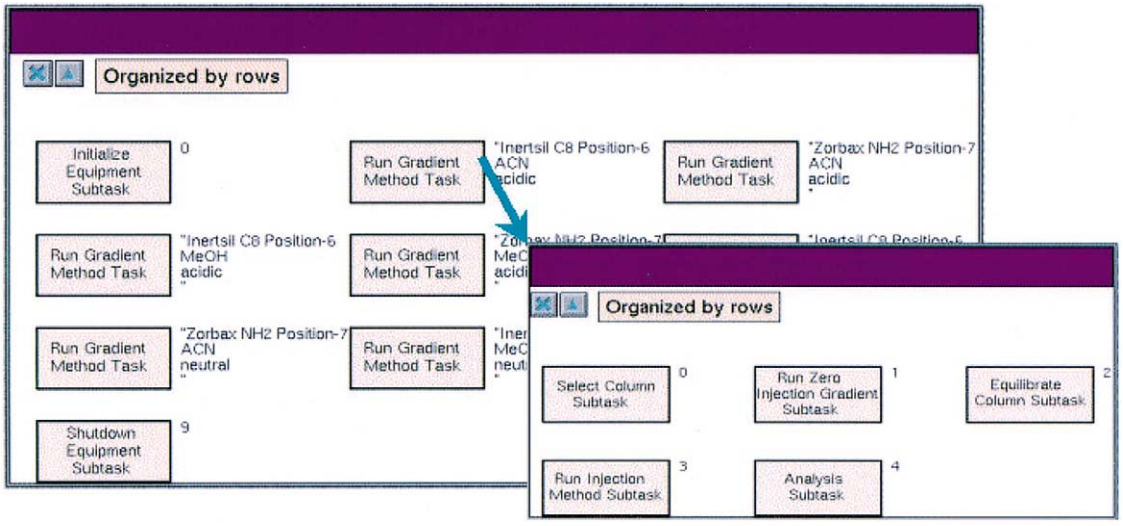


Fig. 2. Tasks and subtasks within a Separation Plan as employed by LabExpert. In this example, the PC computer screen presentation permits visualisation of the various tasks organised as icons in rows. Within each task icon, a series of subtasks is shown.

matographic peak information into a series of ‘‘peak objects’’ for reasoning by decision trees.

3.7. LabExpert design and decision tree operation

Fig. 1 illustrates the high-level logic involved in method development using LabExpert. The user provides inputs on the configuration of the HPLC instrument, which includes comment on the available mobile-phases as well as available columns on the column switching system. The user can also enter any information available about the sample. After the chromatographer defines the criteria for an acceptable separation, the program makes an initial plan, which can later be modified by autonomous decisions after evaluating the information obtained during or at the end of each analysis. As a tool to automate these decisions, LabExpert uses a combination of linear and non-linear chromatographic theory and heuristic rules derived from the knowledge base of experienced chromatographers and imbedded into the software as a series of programmable equation objects and rules. The background, interrelatedness and origin of these theoretical representations, equations and heuristic rules that describe the linear and non-linear chromatographic behaviour of non-polar and polar solutes under different chromatographic modes and elution protocols have been described [46–48] in recent reviews.

3.8. LabExpert program elements and software architecture

When performing method development for a particular sample object, LabExpert creates an ‘‘analytical-session-development session’’ object referred to as the ‘‘ASD’’. The ASD contains the plans that have been executed, those awaiting execution and all of the results from previous chromatographic runs. The ASD is passed to decision trees within LabExpert for reasoning. The ASD includes input objects, chromatographic selectivity objects, a planning object containing the current plan of operations, and results from previous and present HPLC separation methods. When LabExpert performs any type of HPLC run, such as a zero-injection gradient, column re-equilibration or sample injection (isocratic or gradient analysis), LabExpert creates an HPLC-method-record (MR), an object containing a complete record of the HPLC method. All of this parameterised information, which is down-loaded to the Chemstation from LabExpert, is stored in the MR. For example, if an Agilent series 1100 HPLC system is used, the calculated chromatographic peak tables as well as spectral purity results are stored in the MR object. The ASD contains all of the HPLC-method-records for a particular method development session. The overall evaluation criteria are compared to current MR results. When a decision tree receives the ASD, it has all of the information necessary to reason about the method development process and to decide the next step, such as selection of the preferred attributes of the next experiment.

The design features of LabExpert allow the system to run autonomously and perform real-time automation and supervisory control of the HPLC equipment, using the instrument manufacturer’s software to operate the physical functions of networked HPLCs. All of its normal instrument and diode array control functions, data acquisition, and integration functions remain in the system software. The LabExpert system controls the integration parameters and relays event commands to the HPLC workstation, reads the results from the HPLC workstation and then acts upon the results. These results received by the expert system, such as resolution factors, retention times, selectivity ranges and photodiode array spectral data for each peak, are interrogated by LabExpert system and compared to the user-defined acceptability criteria.

3.9. Dynamic planning engine

LabExpert creates an initial plan, based on the user’s preferences and selectivity variables, which contains a sequence of task objects. For example, running a gradient scouting method is represented within the dynamic planning engine as a type of task. Inside of each task object, a series of specific subtasks will be executed. As shown in Fig. 2, a gradient task contains the subtasks of column selection, running a zero-injection gradient, running a column equilibration, running an injection method, and finally running an analysis. Underneath each subtask, there exist a number of decision trees, which provide the detailed knowledge, rules, formulas and equations on how to accomplish the particular subtask. The planning engine is dynamic, as the results

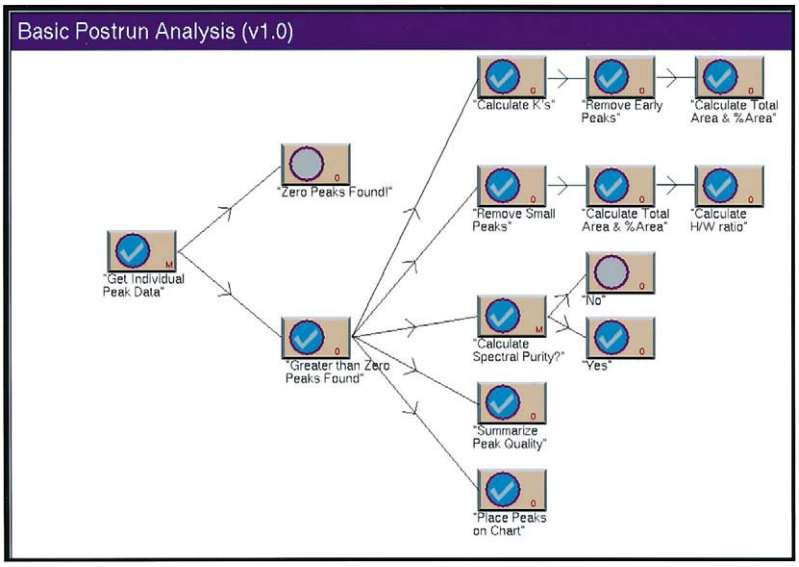


Fig. 3. Illustrative example of the types of decision trees used in the implementation of the expert analytical reasoning approach of LabExpert. In this case, a simple decision tree is shown for performing basic peak analysis. As an action is executed, the decision-tree node changes colour to indicate the reasoning path, and the appropriate notes are added to the audit trail record log.

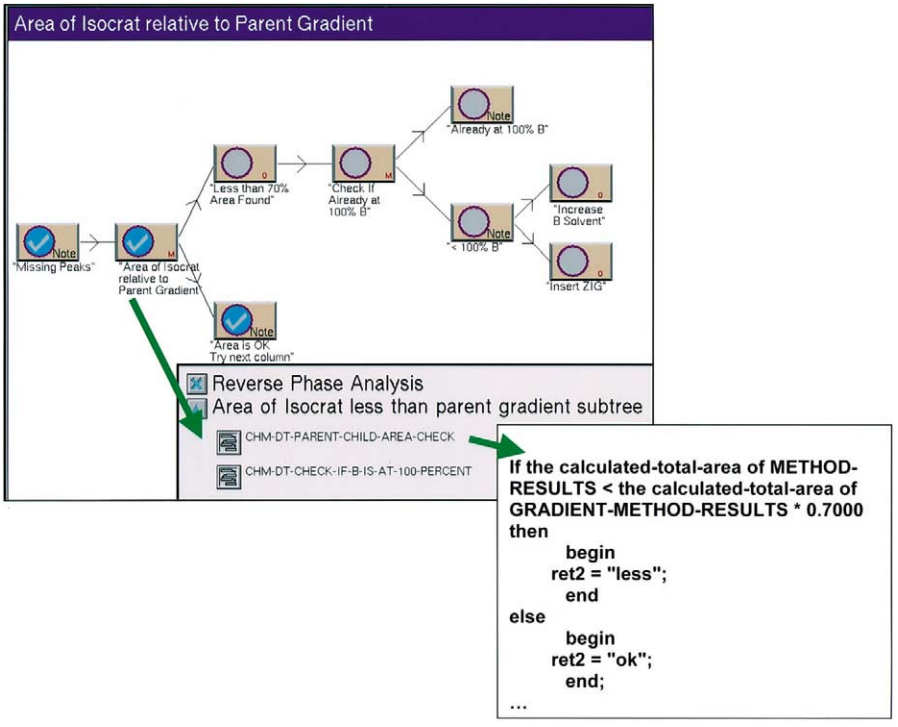


Fig. 4. Run snapshot illustrating the level of programming detail associated with one decision tree node within LabExpert system. In this case, when evaluating the relative total area of an isocratic run to the original scouting gradient, the decision node calls a procedure that implements checks to see if the total isocratic area is less than 70% of the gradient area. These details underneath a single decision tree node can be archived in the usual manner electronically, thus adding to the required level of GMP audit capability with this machine-learning intelligent system.

from the analysis subtask may create a recommendation object, which inserts a new task into the task queue. This may happen, for example, when a combination of solvent, column configuration, sorbent type and mobile phase composition look promising and the system attempts to pursue a line of reasoning which optimises the separation in the isocratic elution mode. Depending on the users' initial choices, the planning engine may stop after finding the first method that meets all of the acceptance criteria, or it may go forward exhaustively trying out all tasks specified in the initial plan.

3.10. Significance of the decision trees

LabExpert uses ''decision trees'' to implement the expert analytical reasoning. A decision tree is a series of actions, incorporating complete object information about the current analytical session as an input. Each node in the decision tree applies expert heuristic rules for chromatography and these rules operate on the objects by adding new information and by creating new objects. When analysing the results of an HPLC method, the decision trees not only apply rules but also extensively apply mathematical and statistical analyses to the chromatographic data. Using the results from these decision trees, LabExpert either plans to run additional methods or reports the results to the user if an HPLC method that meets or exceeds the acceptability criteria is found. LabExpert can then automatically shut down the HPLC system. Decision trees can be displayed, allowing the user to monitor the system status. Fig. 3 shows a simple decision tree, called ''Basic Postrun Analysis''. As LabExpert executes each decision, the node icon changes colour. Results from key decision nodes are added to the log notes of the method record.

Fig. 4 shows the programming detail associated with one decision node within the LabExpert system. There is a subworkspace under each node, involving the procedure to be performed by the program. By encapsulating the expert knowledge into a series of procedures underneath each node, LabExpert is able to efficiently organise, capture and execute the expert knowledge. Each decision tree contains all the information and choices necessary to perform a given action. Fig. 5 illustrates the decision tree for one specific sequence, equilibrating the column, demonstrating the sophistication of that sequence. The entire analytical system is monitored in real-time by a series of decision trees.

LabExpert is designed to not only automate the system but also to direct the current analysis, scrutinise in real-time the results, evaluate the data, make decisions concerning system readiness and to troubleshoot potential instrumental problems. In addition to directing the analysis, LabExpert evaluates the system functions such as column pressure and directs the system to switch to alternate columns or pumps, if necessary. The system also updates decision information continually to the ASD, which can be accessed by an off-site operator to evaluate the status of the experiment or the HPLC system at any time.

3.11. Column preparation and equilibration

As an exemplar of the system, the following experiments illustrate the development of an isocratic separation by LabExpert of the test analytes, CP-547,224, CP-195,543, CP-316,698, and CP-547,238. In this specific case, the Zorbax SB-CN column was selected for illustrative purposes as the initial column, since it was known from other studies that this sorbent was a less than optimal choice for the separation of these compounds. The purpose of this decision entry point was to follow the ability of LabExpert program to determine column unsuitability with these analytes and to then to decide when and how to switch columns. Prior to beginning a sequence of analyses, LabExpert executes an initialisation subtask. This subtask consists of a series of actions to ensure a clean column and stable baseline. Initialisation is begun with a rapid Zero Injection Gradient (ZIG) ranging from a mobile phase of weak elutropicity to a mobile phase of high elutropicity (Fig. 6) to ensure that all previous analytes had been eluted from the column and a stable (reproducible) baseline achieved. A ZIG is repeated after every switch to a different column/sorbent or to a new mobile phase composition. On completion of the ZIG, LabExpert starts an equilibration procedure, without any sample injection, as shown in Fig. 7 using the initial conditions of the next analytical run. As soon as a stable baseline was detected (baseline stability can be defined by the user or by a system default), a decision sequence stops the equilibration

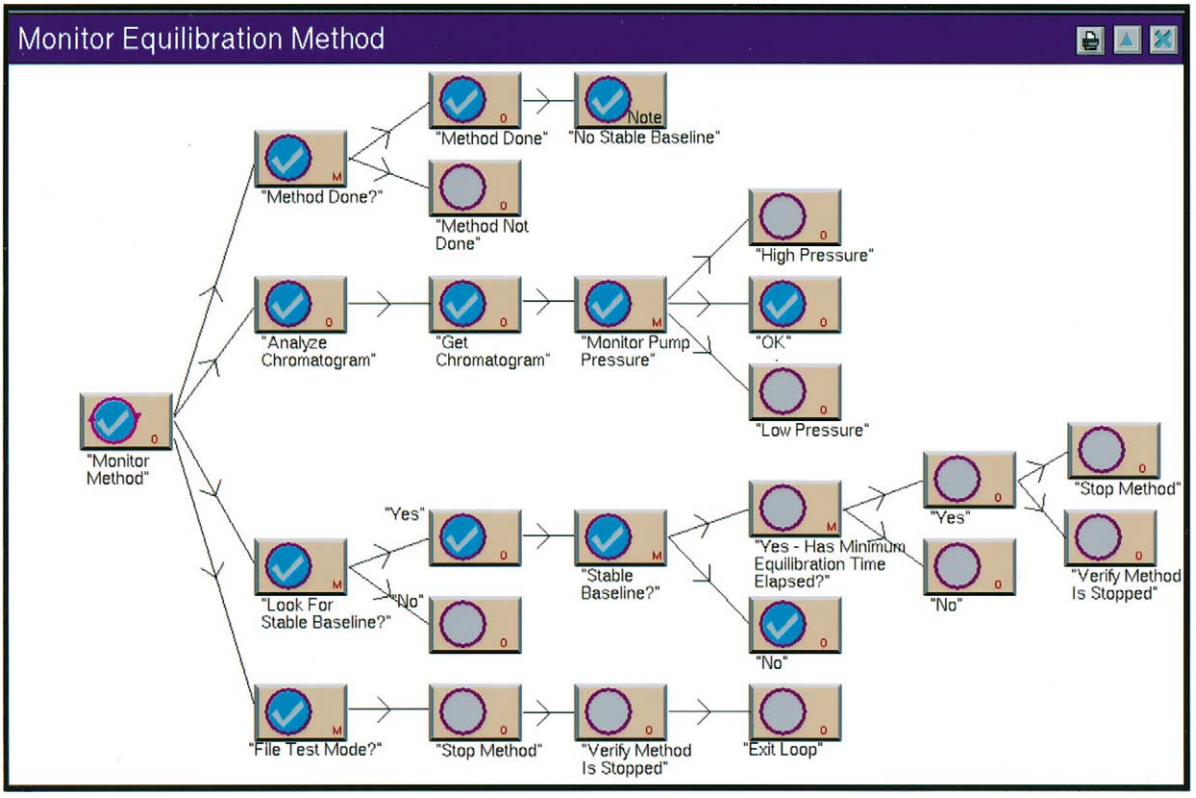


Fig. 5. Decision tree illustrating the LabExpert automatic detection logic for column equilibration. While the HPLC instrument running, LabExpert monitors the diode area UV signal for a stable baseline. When a stable baseline is detected, LabExpert stops the method, effectively ending equilibration early, and then begins an injection method.

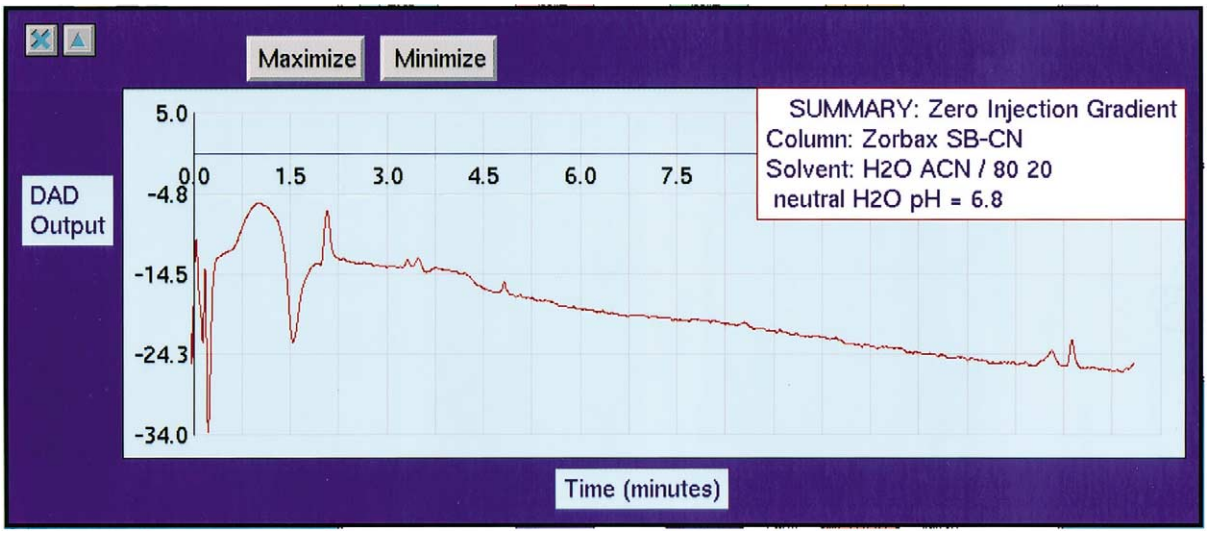


Fig. 6. The initialisation of the column preparation as achieved by the Zero-Injection Gradient (ZIG) with the Zorbax SB-CN column from H_2O, pH 6.8, to ACN–water (20:80, v/v) at pH 6.8 at a flow-rate of 1.0 ml/min and a linear gradient time of 15 min. The purpose of a ZIG is to clean any unknown retained material from the column in preparation for subsequent column equilibration.

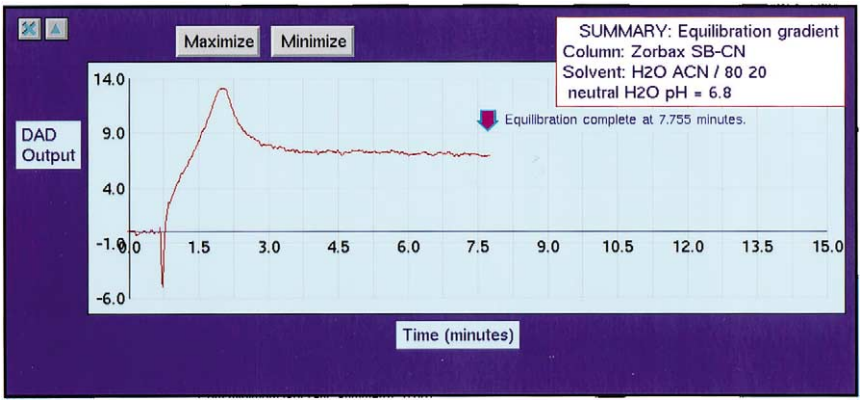


Fig. 7. Initialisation of the column equilibration as a first equilibration experiment following elution of the Zorbax SB-CN column with H_2O, pH 6.8, to ACN–water (20:80, v/v) at pH 6.8 at a flow-rate of 1.0 ml/min. Note how equilibration takes less than 8 min. A chromatographer running this without LabExpert may have pre-programmed the instrument to equilibrate for an hour. In this manner, LabExpert saves time to minimise method development time.

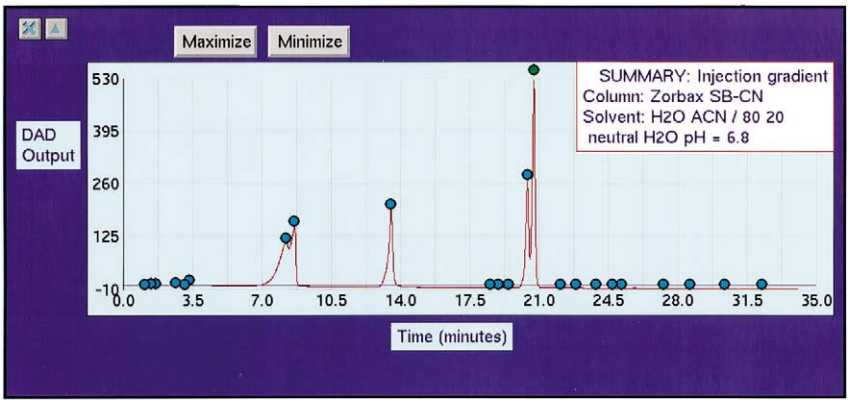


Fig. 8. The first gradient elution experiment with CP-547,224, CP-195,543, CP-316,698 and CP-547,238 separated on the Zorbax SB-CN column, from H_2O, pH 6.8, to ACN–water (20:80, v/v) at pH 6.8 at a flow-rate of 1.0 ml/min and a linear gradient time of 35 min. Concomitant with the record of the detector signal, in real time the LabExpert calculates and appraises the resolution, peak efficiency, analysis time, selectivity, etc., against the user-generated acceptability criteria. Specific node points, activated as cursor-driven sampling points for the calculation of the experimental results against such resolution acceptability criteria, are highlighted as circles.

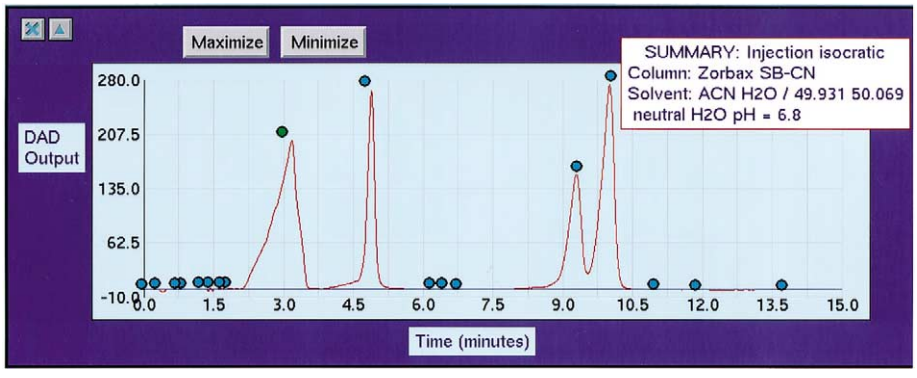


Fig. 9. The isocratic elution profile obtained with CP-547,224, CP-195,543, CP-316,698 and CP-547,238 separated on the Zorbax SB-CN column eluted with ACN–water (49.9:50.1, v/v) at pH 6.8, obtained as the first machine-learning experiment.

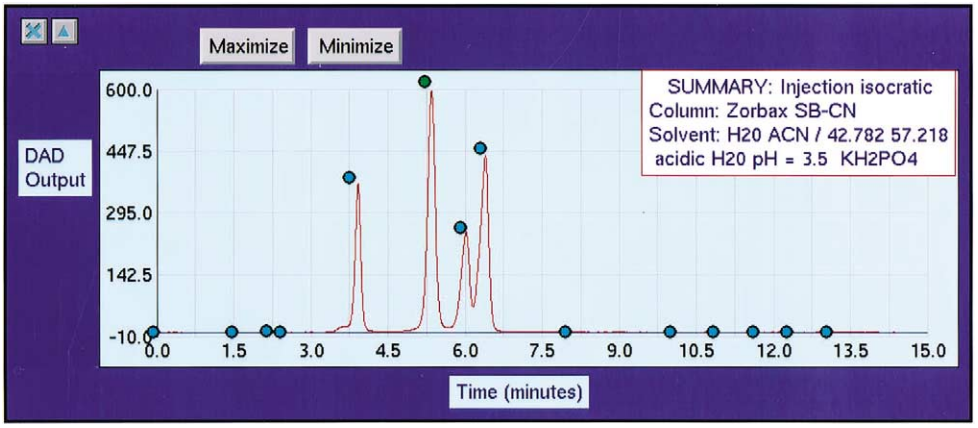


Fig. 10. The isocratic elution profile obtained with CP-547,224, CP-195,543, CP-316,698 and CP-547,238 separated on the Zorbax SB-CN column with ACN–water (42.8:57.2, v/v), and 25 m*M* KH_2PO4, pH 3.5, as the second machine-learning experiment.

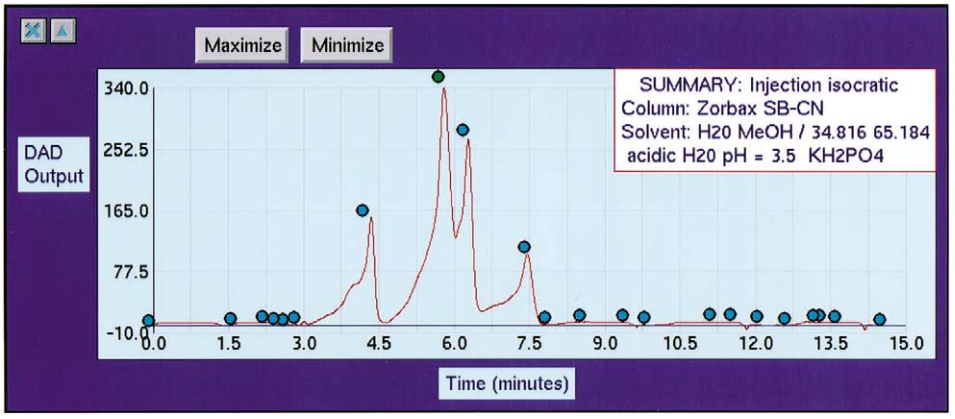


Fig. 11. The isocratic elution profile obtained with CP-547,224, CP-195,543, CP-316,698 and CP-547,238 separated on the Zorbax SB-CN column with MeOH–water (65.2:34.8, v/v) and 25 m*M* KH_2PO4, pH 3.5, as the third machine-learning experiment.

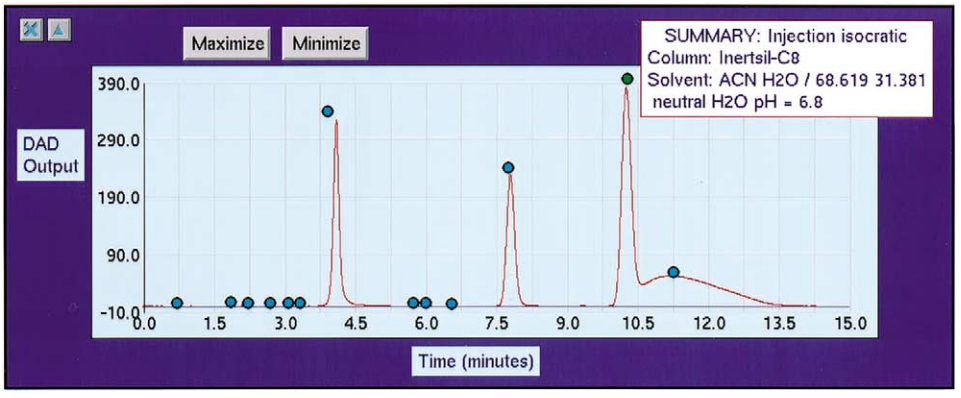


Fig. 12. Isocratic elution profile obtained with CP-547,224, CP-195,543, CP-316,698 and CP-547,238 separated on the Inertsil-C_8 column with ACN–water (68.6:31.4, v/v) at pH 6.8, as the fourth machine-learning experiment.

procedure and LabExpert then proceeds to the next stages of the analysis.

3.12. Selection of gradient screening conditions and elution conditions for the isocratic runs on initial column

The initial analytical run for this particular analysis involved an acetonitrile–water gradient, as shown in Fig. 8. Completion of the gradient triggered a sequence of decisions, based on the initial chromatographic results, to determine the composition of the initial isocratic mobile-phase composition that was expected to achieve partial resolution of CP-547,224, CP-195,543, CP-316,698 and CP-547,238. The initial isocratic analysis so selected by LabExpert was a acetonitrile–water (50:50, v/v) eluent on the Zorbax SB-CN column. As is apparent from Fig. 9, these conditions yielded a chromatogram of poor peak shape and insufficient resolution between two of the peaks. Depending on the compositions of the buffers within the HPLC system reservoirs, a feature of the LabExpert system is that pH ''scouting'' can be employed in a manner similar to that exploited for solvent strength ''scouting'' with organic solvents. Thus, based on this knowledge and/or knowledge retrieved from the compound descriptor file, LabExpert automatically decided without operator intervention to change to a more acidic mobile phase. The sample was reinjected using the same column with a water–acetonitrile (43:57, v/v) phase adjusted to pH 3.5 (Fig. 10), after an equilibration run to determine baseline stability. Although the peak shape was improved, there was still insufficient retention and the resolution of the third and fourth peaks was still inadequate accordingly to previously set user-defined acceptability criteria. Because subsequent analysis with lower acetonitrile mole fraction content still did not provide sufficient resolution, LabExpert then automatically changed the organic modifier of the mobile phase to methanol. As indicated in Fig. 11, this change did not improve the quality of the analysis and, in fact, decreased resolution and compromised other peak characteristics.

3.13. Screening additional columns

At this point, the reasoning processes within LabExpert determined that the initial column was not suitable for this separation. The column was automatically switched firstly to a Waters Symmetry C_{18} and then to an Inertsil-C_8 column with the Zero Injection Gradient sequence repeated in each case. Following the procedures outlined above for the initial column, LabExpert arrived at a separation that matched the acceptability criteria for the test compounds, which in this case involved the use of the Inertsil-C_8 column. Clearly, with other types of analytes the LabExpert's decision outcomes could lead to the identification of other types of columns/sorbents as the preferred choice(s) from a panel of columns/sorbents of different separation selectivity. Such outcomes have been readily documented in these laboratories with various, more complex mixtures of polar analytes (I. Ting-Po, S. Guhan, K. Taksen, D. Myers, M.T.W. Hearn, 2002, unpublished results) using the various capabilities of LabExpert. With the LC Spiderling 8 position switching valve system used in these studies, up to eight different columns can be consecutively evaluated, although in principle there is no upper limit on the number of columns that can be screened using LabExpert. Specific decisions concerning changes in mobile phase composition were based on actual chromatographic results from the different columns. Fig. 12 illustrates another chromatogram obtained in this machine-learnt sequence with the Inertsil-C_8 column. The final separation, as shown in Fig. 13, displays excellent resolution and peak shape.

3.14. System shutdown

When LabExpert program has determined that an acceptable separation had been achieved according to the match with the pre-defined acceptability criteria, a shutdown sequence is automatically initiated. The column is eluted with a suitable solvent mixture to achieve equilibrium and the instrument returned to a standby state by turning off the pump and optionally the lamp for the diode array detector. In the event of instrument malfunction, or when no conditions can be generated that approach the acceptability criteria set initially by the user/operator, LabExpert shuts the system down. A full technical report with graphics is then generated, detailing the sequence of analyses and decisions. With LabExpert every analy-

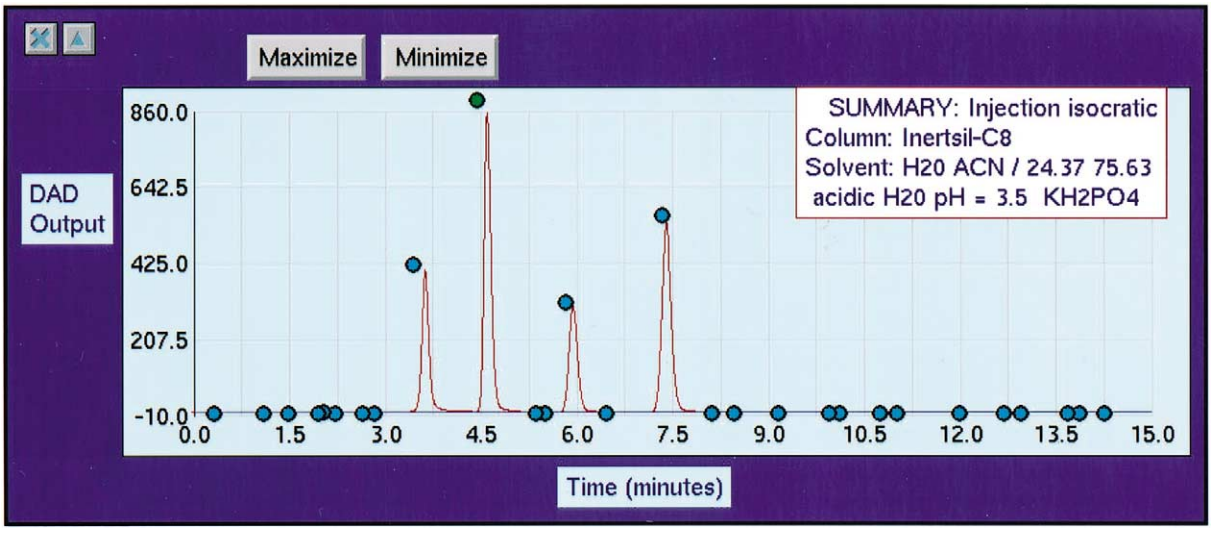
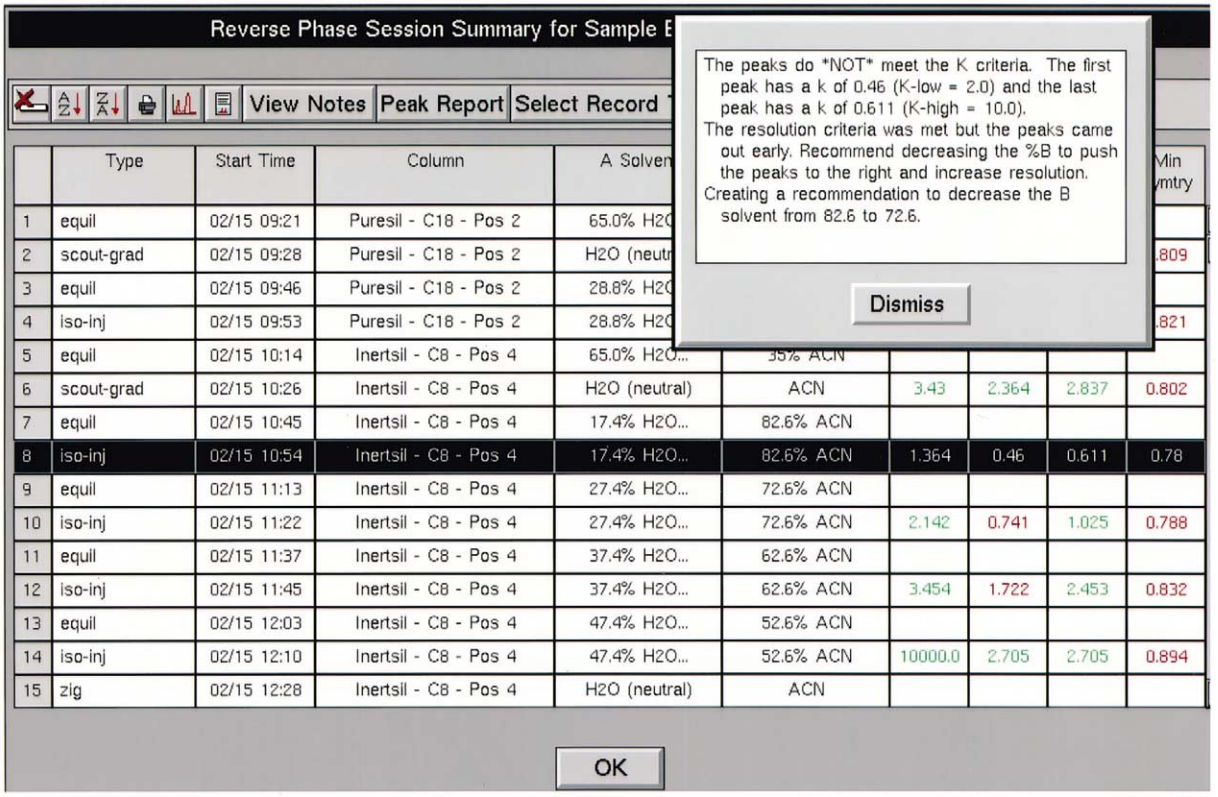


Fig. 13. Isocratic elution profile obtained with CP-547,224, CP-195,543, CP-316,698 and CP-547,238 separated on the Inertsil-C_8 column with ACN–water (75.6:24.4, v/v) in 25 m*M* KH_2PO4, pH 3.5, with attainment of resolution that matched the acceptability criteria.



	Type	Start Time	Column	A Solven					Min ymtry
1	equil	02/15 09:21	Puresil - C18 - Pos 2	65.0% H2O					
2	scout-grad	02/15 09:28	Puresil - C18 - Pos 2	H2O (neutr					.809
3	equil	02/15 09:46	Puresil - C18 - Pos 2	28.8% H2O					
4	iso-inj	02/15 09:53	Puresil - C18 - Pos 2	28.8% H2O					.821
5	equil	02/15 10:14	Inertsil - C8 - Pos 4	65.0% H2O...	35% ACN				
6	scout-grad	02/15 10:26	Inertsil - C8 - Pos 4	H2O (neutral)	ACN	3.43	2.364	2.837	0.802
7	equil	02/15 10:45	Inertsil - C8 - Pos 4	17.4% H2O...	82.6% ACN				
8	iso-inj	02/15 10:54	Inertsil - C8 - Pos 4	17.4% H2O...	82.6% ACN	1.364	0.46	0.611	0.78
9	equil	02/15 11:13	Inertsil - C8 - Pos 4	27.4% H2O...	72.6% ACN				
10	iso-inj	02/15 11:22	Inertsil - C8 - Pos 4	27.4% H2O...	72.6% ACN	2.142	0.741	1.025	0.788
11	equil	02/15 11:37	Inertsil - C8 - Pos 4	37.4% H2O...	62.6% ACN				
12	iso-inj	02/15 11:45	Inertsil - C8 - Pos 4	37.4% H2O...	62.6% ACN	3.454	1.722	2.453	0.832
13	equil	02/15 12:03	Inertsil - C8 - Pos 4	47.4% H2O...	52.6% ACN				
14	iso-inj	02/15 12:10	Inertsil - C8 - Pos 4	47.4% H2O...	52.6% ACN	10000.0	2.705	2.705	0.894
15	zig	02/15 12:28	Inertsil - C8 - Pos 4	H2O (neutral)	ACN				

Fig. 14. Illustrative example of the format of the LabExpert spreadsheet interface for viewing the results of an analytical-session-development (ASD) session and reporting for the purposes of an audit trail as set out in GMP/GLP requirements.

sis and decision step of the process is recorded electronically, providing a database for GMP audit tracing. All the information obtained, from unacceptable separations, as well as the final separation meeting the acceptability criteria can be incorporated into the decision process for future methods development. If LabExpert cannot obtain an acceptable method, it rates each attempt on a normalised rank-order basis. The investigator can then review those experiments to determine whether any of the separations meet revised criteria or whether a new experimental sequence is required. As part of its investigator interface, LabExpert provides a spreadsheet (cf. Fig. 14), which summarises the results of an entire method development session. From this spreadsheet, the investigator can view all of the intermediate and final results including the decision notes, various chromatograms, and peak reports. The investigator can also quickly review the method development session in a electronic slide show format.

4. Conclusions

In the case of the exemplar described above, the above set of experiments with the four pharmaceutical compounds was started at 15:30 h and ran unattended in an operator-free environment until a satisfactory separation was achieved. Initial conditions were (purposefully) chosen to start with an ''unsatisfactory'' column in order to demonstrate the knowledge-acquisition capabilities of the software. LabExpert identified a satisfactory separation at 04:30 h, after an unattended 13-h experimental time course and then shut down the system. The experiments included six Zero Injection Gradient (ZIG) runs, 18 system equilibration runs, six gradient elution analyses, and 12 isocratic elution analyses with three columns packed with three types of adsorbent (Zorbax SB-CN, Waters Symmetry C_{18} and Inertsil C_8).

As part of its software architecture, LabExpert constantly monitors the HPLC instrumentation and makes decisions based on the current status of the equipment and associated system features. This allows the system itself to determine that the column and elution profile baseline are stable, that the mobile phase supply and composition are appropriate for the planned studies and that reequilibration has been achieved, prior to proceeding to the next injection. This feature of LabExpert introduces significant timesaving, since the alternative would be to program a very long equilibration time to ensure that the system always reaches a stable condition. LabExpert monitors also the HPLC system for fault conditions, such as low, high or fluctuating column pressure; mobile phase leaks; faulty injections or exhausted solvent supply. Should a fault be detected, then the fault is corrected, e.g., by automatically switching to alternate solvent reservoirs or if no corrective action is available, LabExpert initiates a shutdown sequence. If it detects a high back-pressure problem, LabExpert automatically reduces the flow-rate in an attempt to reduce the pressure.

The separation outlined above is an actual method development process, using LabExpert as installed at Pfizer's Groton CT laboratory. LabExpert has been successfully applied in numerous other studies, including the separation of chiral analytes on chiral sorbents and basic and hydrophobic substances and biomolecules by RP-HPLC, running with minimal user intervention or involvement. These LabExpert applications have demonstrated the unique ability to not only automate the execution of the HPLC methods but also to analyse the results in real time. Because LabExpert uses a knowledge-driven search for the best solution within the total separation space of a particular analytical problem, it provides a way to standardise methods and to use previous machine-learnt experience. The program is a dynamic self-monitoring system, capable of learning as new information is acquired or added. Because LabExpert actively plans each method step and analyses the results during or at the conclusion of each procedure, the experiment often ends with attainment of an acceptable method in a time frame shorter than obtained with current empirical procedures, with better utilisation of laboratory resources and enhanced performance by personnel. HPLC expertise and system monitoring is available and automated 24 h a day, 7 days a week. Highly trained staff with expertise in chromatography and analytical chemistry are thus relieved of many routine tasks, allowing them to focus on more difficult separations that require more critical input. Since default decision

modes, based on expert experience and documented heuristic rules, have been incorporated into LabExpert system analysts with less experience are thus able to access and efficiently use this knowledge within laboratories at alternative locations where immediately access to the experienced chromatographic analysts is precluded due to distance constraints.

References

[1] A.M. Siouffi, R. Phan-Tan-Luu, J. Chromatogr. A 892 (2000) 75.
[2] J.L. Glajch, L.R. Snyder, J. Chromatogr. 485 (1989) 1, Special Issue: Computer-Assisted Method Development in Chromatography.
[3] P. Hubert, P. Chiap, M. Moors, B. Bouruigignon, D.L. Massart, J. Crommen, J. Chromatogr., A 665 (1994) 87.
[4] Y. Hu, D.L. Massart, J. Chromatogr. 485 (1989) 311.
[5] J. Fulper, J. Chromatogr. 488 (1989) 597.
[6] D. Wu, M. Berna, G. Mair, J. Johnson, J. Pharm. Biomed. Anal. 16 (1998) 57.
[7] B. Olsen, P. Tang, Process Control Qual. 10 (1997) 25.
[8] R. Dappen, I. Molnar, J. Chromatogr. 592 (1992) 133.
[9] P. Zhu, J. Dolan, L. Snyder, J. Chromatogr. A 756 (1996) 41.
[10] J. Dolan, L. Snyder, R. Wolcott, P. Haber, T. Bacczek, R. Kaliszan, L. Sander, J. Chromatogr. A 857 (1999) 41.
[11] Q. Zhu, M. Stillman, J. Chem. Inf. Comput. Sci. 36 (1996) 497.
[12] S.D. Patterson, J. Chromatogr. 592 (1992) 43.
[13] P. Csokan, F. Darvas, F. Csizmadia, K. Valko, LC·GC 12 (1994) 1.
[14] J. Fekete, G. Morovjan, F. Csizmadia, F. Darvas, J. Chromatogr. A 660 (1994) 33.
[15] S.V. Galushko, K. Deguchi, M. Ito, Chromatography (Japan) 18 (1997) 298.
[16] J.L. Martinez-Vidal, P. Parrilla, A.R. Fernandez-Alba, R. Carreno, F. Herrera, J. Liq. Chromatogr. 18 (1995) 2969.
[17] A.G. Wright, A.F. Fell, J.C. Berridge, Chromatographia 24 (1987) 533.
[18] R. Faulstich, T. Catalano, LC·GC 9 (1991) 11.
[19] R. Cela, M. Lores, Comp. Chem. 20 (1996) 175.
[20] Getting Started with G2 Tutorials Version 5.0, Gensym Corporation, Cambridge, MA, 1997, pp. 1–35.
[21] M. Peris, Crit. Rev. Anal. Chem. 26 (1996) 219.
[22] J. Namiesnik, Crit. Rev. Anal. Chem. 30 (2000) 221.
[23] R. Kurzweil, in: The Age of Intelligent Machines, Massachusetts Institute of Technology, 1990.
[24] M. Mulholland, D.B. Hibbert, P.R. Haddad, P. Parslov, Chem. Intell. Lab. Systems 30 (1995) 117.
[25] M.J. Stillman, Expert systems-analytical applications, in: A. Townshend (Ed.), Encyclopedia of Analytical Science, Vol. 2, Academic Press, London, 1995, p. 1233.
[26] N. Ulrich, G. Wunsch, J. Prakt. Chemie-Chemiker-Zeitung 337 (1995) 628.
[27] M. Mulholland, K. Mckinnon, P.R. Haddad, J. Chromatogr. A 739 (1996) 25.
[28] X. Rius F, Anal. Chim. Acta 283 (1993) 518.
[29] H. Du, M.J. Stillman, Anal. Chim. Acta 354 (1997) 77.
[30] S. Lahiri, M.J. Stillman, Anal. Chem. 64 (1992) 283.
[31] L.J. Hu, E.F. Saulinskas, P. Johnson, P.D. Harrington, Comp. Appl. Biosci. 12 (1996) 311.
[32] C.H. Bryant, A. Adam, D.R. Taylor, R.C. Rowe, Anal. Chim. Acta 297 (1994) 317.
[33] N. Ulrich, G. Wunsch, J. Prak. Chem. Chem. Z. 339 (1997) 441.
[34] J.O. Debeer, C.V. Vandenbroucke, D.L. Massart, B.M. Despiegeleer, J. Pharm. Biomed. Anal. 14 (1996) 525.
[35] A. Tsantilikakoulidou, I. Panderi, F. Csizmadia, F. Darvas, J. Pharm. Sci. 86 (1997) 1173.
[36] Z. Ramadan, M. Mulholland, D.B. Hibbert, Chem. Intel. Lab. Syst. 40 (1998) 165.
[37] H. Du, G.S. Huang, M.J. Stillman, Anal. Chim. Acta 324 (1996) 85.
[38] Z. Ramadan, M. Mulholland, D.B. Hibbert, P. Preston, P. Compton, P.R. Haddad, J. Chromatogr. A 804 (1998) 29.
[39] C.H. Bryant, A.E. Adam, Chem. Intell. Lab. Syst. 34 (1996) 21–40.
[40] L.R. Snyder, J.J. Kirkland, J.L. Glajch, in: Practical HPLC Method Development, 2nd ed., Wiley, New York, 1997.
[41] A.A. Agbodjan, S.C. Rutan, Chem. Intel. Lab. Systems 24 (1994) 137.
[42] A.V. Pirogov, M.M. Platonov, I.V. Pletnev, O.N. Obrezkov, O.A. Shpigun, Anal. Chim. Acta 369 (1998) 47.
[43] M.C. Bruzzoniti, E. Mentasti, Sarzanini, J. Chromatogr. B 717 (1998) 3.
[44] E.I. Klein, S. Rivera, J.E. Porter, Biotechnol. Prog. 16 (2000) 506.
[45] S.A. Wood, J.M. Long, R.J. Simmonds, J.W. Bridges, D. Stevenson, J. Pharm. Biomed. Anal. 16 (1997) 231.
[46] M.T.W. Hearn, in: S. Ahuja (Ed.), Handbook of Bioseparation, Academic Press, San Diego, CA, 2000, p. 72.
[47] M.T.W. Hearn, in: K. Gooding, F.E. Regnier (Eds.), HPLC of Biopolymers, Marcel Dekker, New York, 2001, p. 99.
[48] R.I. Boysen, M.T.W. Hearn, In: R. Simpson (Ed.), Chromatography Protocols for Protein Analysis: A Laboratory Manual Cold Spring Harbour Laboratory Press, NY, 2002, in press.